BAuA
Federal Institute for Occupational Safety and Health
Federal Office for Chemicals
Friedrich-Henkel-Weg 1-25 44149
Dortmund, Germany

RE: CLH Report for Glyphosate, EC Number 213-997-4

Dear Sirs,

Below are my comments on the evaluation of carcinogenicity in the CLH Report for Glyphosate (the Report), EC Number 213-997-4, prepared by the Federal Institute for Occupational Safety and Health (BAuA). In my comments, you will see that I disagree with the conclusions on the human epidemiological data and that I find serious flaws in the evaluation of the animal carcinogenicity data. I have also prepared a pooled analysis of the animal carcinogenicity data that clearly indicates the hemangiosarcomas and malignant lymphomas show statistically significant trends even when excluding doses above 1000 mg/kg/day.

I am also including several supplemental files with this submission including all cited papers, the computer code I used to produce the pooled analysis, and the computer code I used to calculate statistical significance for testing the observed data sets against the historical controls. I have also included a manuscript by Ghisi et al. (2016) that does a meta-analysis on the ability of glyphosate to induce micronuclei.

What I found most disturbing with this submission is that, despite our previous concerns about the EFSA conclusions on carcinogenicity, the review continues to disregard guidance set forth by ECHA, OECD, IARC and others on how to evaluate carcinogenicity data, especially regarding the use of the *limited evidence* category for the human data, the appropriate use of historical controls and the proper use of findings of a positive trend in an animal cancer study.

In my opinion, having reviewed a large number of compounds for carcinogenicity and having read both the Report and the ECHA Guidelines, glyphosate should be classified into Group 1b.

Sincerely,

Prof. Christopher J. Portier
Thun, Switzerland
cportier@mac.com
+41 79 605 79 58
July 8, 2016

**Human Evidence**

On page 93 of the Report, the human evidence regarding glyphosate carcinogenicity is summarized as follows:

"*Epidemiological studies revealed partly contradictory results. However, in most studies, no association with an exposure to glyphosate could be established. In particular, the largest study, i.e., the AHS (see above), was negative. Taken together, the epidemiological data does not provide convincing evidence that glyphosate exposure in humans might be related to any cancer type. Epidemiological studies are of limited value for detecting the carcinogenic potential of an active substance in plant protection products since humans are never exposed to a single compound alone. Thus, the results of the studies are associated to different formulations containing glyphosate or mixtures of different active substances.*"

The first sentence claims the results are contradictory. This is only true if classify each study is classified as significant or non-significant. Examining the numerical findings presents a different picture. Table 1 lists the 8 studies (of sufficient quality to be utilized) that evaluated the relationship between non-Hodgkin lymphoma (NHL) and exposure to glyphosate. Simply looking to see if the studies tend to have a relative risk above or below 1 shows the studies to be consistently positive across the board with the exception of the AHS exposure-response analysis (that had problems with classifying the exposure) and the Orsi *et al* study (that had a relative risk of exactly 1). This is quite clearly illustrated using the tree plot in Figure 1.

The sentence '*Taken together, the epidemiological data does not provide convincing evidence that glyphosate exposure in humans might be related to any cancer type.*' is difficult to accept given that the three meta-analyses, all including the AHS study, show a statistically significant association between use of glyphosate pesticides and NHL in humans (Table 2). Finally, the statement that "*the results of the studies are*

*associated to different formulations containing glyphosate or mixtures of different active substances.*" is not supported by actual data so this is speculation and not fact.

In "Guidance on the application of the CLP criteria – Version 4.1", Annex I: 3.6.2.2.3 states that "*The terms 'sufficient' and 'limited' have been used here as they have been defined by the International Agency for Research on Cancer (IARC) and read as follows: ... limited evidence of carcinogenicity: a positive association has been observed between exposure to the agent and cancer for which a causal interpretation is considered to be credible, but chance, bias or confounding could not be ruled out with reasonable confidence.*" The meta-analyses indicate that a positive association has been observed so the only reason you would have for not classifying the human evidence as limited is that you believe the causal relationship is not credible or that the bias and/or confounding is so bad as to make these studies worthless. This is clearly not the case. It is likely that the decision is being skewed by placing too much emphasis on the AHS study; the meta-analysis is designed to avoid this problem.

Finally, this paragraph also implies that human epidemiology data will never be of importance in evaluating a pesticide because the pure compound is not used on humans. Such a statement is not scientifically sound and fails to use the science to address the safety of the public.

**Mouse Carcinogenicity Data**

Also on page 93 of the Report, the data on the carcinogenicity of glyphosate in mice is summarized.

"*In the mouse, the incidences in malignant lymphoma, in renal tumours and haemangiosarcoma in male animals were considered in detail. Slightly higher incidences when compared with concurrent controls were confined to very high dose levels above the OECD-recommended limit dose of 1000 mg/kg bw/day and exceeding the MTD. In addition, the outcome of statistical tests was contradictory. Mostly, but not always, trend tests revealed statistical significance but pairwise comparisons failed to detect a significant difference relative to the control group. The reported incidences of all three tumour types fell within their historical control range which were, however, of variable reliability. If the four studies in CD-1 mice are considered together, it becomes apparent that all tumours were observed also in the control groups and in some groups receiving lower doses in at least one concurrent study. Furthermore, the results were not consistent with regard to dose responses. To conclude, there is not enough evidence to consider the tumours in mice as treatment-related.*"

It is unusual to have four studies in the same species and strain for an evaluation. It is possible to make direct comparisons between the studies and even pool the data for a combined analysis. Table 3 quickly summarizes the findings from the four studies in CD-1 mice and the one study in Swiss mice. One thing that stands out in Table 3 is that the studies were conducted for either 18 months or 24 months. This is a critical difference that does not get much discussion in the Report.

Cancer increases in risk generally as a power of length of exposure (Portier, Hedges and Hoel, 1986). This relationship was used to develop a means to adjust the length of time an animal is on a study, enabling a scientist to determine risk at the end of two-years, the typical time used for animal bioassays (Bailer and Portier (1988) and Portier and Bailer (1988)). This is called the Poly-3 adjustment. The US National Toxicology Program uses the Poly-3 test to evaluate significance in their animal bioassays. Now you will note that three of the mouse studies were only conducted for 18 months. (Comparing 18 month studies with 24 month studies without making an adjustment for the differences in length of exposure is like comparing cancer rates in 40 year-olds exposed for 25 years to cancer rates in 65 year-olds exposed for 50 years and concluding they are not consistent with each other; the conclusion is meaningless because the correct evaluation was not done.) Thus, in order to compare all 5 studies, we must use the Poly-3 adjustment to extrapolate the 18 month studies to estimate what we think the cancer risk would have looked like at 24 months. The adjustment decreases the number of animals without tumors in all groups by $(18/24)^{3}$. The p-values for both the unadjusted trend test and the poly-3 adjusted trend test are given in Table 4 for male mouse renal tumors.

As an example of how the Poly-3 adjustments work, consider a comparison of the high-dose renal tumor response in the 1983 study (3/50=6%) to the high-dose response in the 1997 study (2/50=4%). In the 1997 study, 48 animals had no tumors at 18 months; the poly-3 adjustment reduces this to 20.25 leading to an incidence estimate of 2/22.25=9%. Because the Poly3 test effectively reduces the number of animals on study, even though the incidence estimate goes up, the p-value for the trend test could go down. Numerous evaluations of the validity of the poly-3 adjustment have been published in the peer-reviewed literature and it seems to work very well.

Now that the lengths of the studies have been adjusted, the next question to ask is whether this dose-response is consistent across all of the studies or whether there are anomalies. Combining all of the studies into one pooled analysis (Table 5, Line 1) and performing a trend analysis on the pooled data yields highly significant findings (Table 5, Line 1). Excluding the Swiss Albino mouse study (2001) and only using the CD-1 mice also yields a significant trend (Table 5, Line 3). Repeating these analyses with the Poly-3 adjusted data does not alter the significant findings. Poly-3 adjusted dose-response for renal tumors in the entire set of mouse studies is shown in Figure 2. Here, each dose-response point from each study is plotted along with the 95% confidence bound around the response. It is somewhat hard to see that there is a pattern here that is consistent. To make it easier to see, I pooled all the controls into one group, pooled the animals given doses between 0<dose≤300 in a second group, and similarly for animals given doses between 300<dose≤1500 and dose>1500. These results are plotted against the mean dose in each set of pooled doses in Figure 3 (the horizontal blue lines show the range of the doses that were combined). The trend in the data is more evident in Figure 3 than in Figure 2. The pooled data sets were also analyzed by the unadjusted and poly-3 adjusted trend

tests and shown to be significant (Table 5, Lines 2 and 4). Finally, as noted in the Report, it seems that all of the response is in doses above 1000 mg/kg/day. After removing all doses above 1000 mg/kg/day and repeating all of the analyses, the results of the analysis are shown in Table 5, Lines 5-8. Without the doses above 1000 mg/kg/day, the effect disappears.

Tables 6 and 7 repeat these analyses for malignant lymphomas and Figures 4, 5, and 6 show the resulting plots of the data. In Figure 4, it is easily seen that the Swiss mice had a very different background tumor rate compared to the CD-1 mice so for the remaining two Figures (5 and 6), only CD-1 mice are plotted. Because of the different backgrounds between the Swiss mice and the CD-1 mice, when they are all combined, the joint analysis is not significant (Table 7, lines 1 and 2). Removing the Swiss mouse study and only evaluating the CD-1 mice leads to highly significant trends in all analyses (Table 7, lines 3-8). A significant trend remains even after removing the doses>1000 (Table 7, lines 5-8) suggesting this is not a high-dose only effect. This is very clear when you examine Figure 7.

Tables 8 and 9 repeat these analyses for hemangiosarcomas and Figures 7 and 8 show the resulting plots of the data. The findings in the Swiss mouse were unclear in the reporting so these tables only contain analyses of the CD-1 mouse data. All analyses are highly significant (Table 9) and they remain significant if doses>1000 are excluded (Table 9, lines 3 and 4). So again, this is not a high dose-only effect.

With these analyses, certain things are clear. The statement "*If the four studies in CD-1 mice are considered together, it becomes apparent that all tumours were observed also in the control groups and in some groups receiving lower doses in at least one concurrent study."* is highly misleading. Combining all four studies in CD-1 mice leads to very strong statistical significance in the data. Also, "*Furthermore, the results were not consistent with regard to dose responses.*" is also incorrect and not actually supported by the data. Finally, the statement "*Slightly higher incidences when compared with concurrent controls were confined to very high dose levels above the OECD-recommended limit dose of 1000 mg/kg bw/day and exceeding the MTD."* while partially correct is also very misleading. When doses above 1000 mg/kg/day are excluded, the pooled data from the four CD-1 mouse studies remain significant for both the malignant lymphomas and the hemangiosarcomas. Also, the OECD-recommended limit is not the MTD (maximum tolerated dose) and showing exceedance of an MTD requires more information than simply that the dose was large.

Given a careful, objective evaluation of these data, I strongly suggest you change your conclusion from the mouse studies from "*To conclude, there is not enough evidence to consider the tumours in mice as treatment-related.*" to "***To conclude, there is enough evidence to consider the tumours in mice as treatment-related***."

Finally, a few comments on the reviews of the individual studies starting on page 67 of the Report.

Page 68 - "*Obviously, the carcinogenicity study in Swiss albino mice by Kumar (2001, ASB2012-11491) revealed an increase in malignant lymphoma incidence over the control at the top dose level of around 1460 mg/kg bw/day in both sexes but the background (control) incidence was also quite high. In fact, at least in males, the number of affected animals in the control groups was markedly higher in this strain than in three studies in CD-1 mice. It must be emphasised that this tumour is quite common in ageing mice and that Swiss mice are frequently affected (for details, see below). In this study, malignant lymphoma accounted for 54.6% of the total number of tumours when all groups are considered together.*" Without actually using historical controls, an attempt is made here to downplay the significance of this finding by saying the concurrent control was high. And then it is not clear at all why the 54.6% figure is put into this paragraph. Is this study positive? Yes. Are there flaws in this study? No. Why does this Report then downplay this finding? Especially when you see similar findings in the other studies?

Page 68 -"*In the most recent study in CD-1 mice by Wood et al. (2009, ASB2012-11490), there was a higher incidence of the same tumour type in high dose males (5/51 vs. 0/51 in the control group). Likewise, in the study by Sugimoto (1997, ASB2012-11493), there were a higher number of male mice affected at the exaggerated dose level of 40000 ppm (approx. 4350 mg/kg bw/day) than in the control group (6/50 vs. 2/50). In the study by Atkinson et al. (1993, TOX9552382), in contrast, there was no dose response and the incidence in the control group was similar to that at the top dose level.*" Regardless, this entire paragraph is attempting to compare control animals ranging over 16 years with differing terminal sacrifice times and from different laboratories. Such a comparison is inappropriate because of the known drift in strains over time and increasing tumor risk with age. The OECD guidelines make this very clear.

Page 69 – "*The trend test also provided a p-value above the significance level of 0.05, most probably because of the high control incidence (see Table 33).*" The p-value for trend is 0.0535122, technically above 0.05, but it is misleading when trying to compare across studies not to mention that this is almost significant.

Page 69 - "*In contrast, re-analysis of the studies by Wood et al. (2009, ASB2012-11490) and Sugimoto (1997, ASB2012-11493) showed statistically significant increases with dose for male CD-1 mice in the trend test (Table 34 and Table 35) but a rather low or even "zero" incidence in the control groups might be behind this finding.*" Where are the historical controls to support the speculation in the last part of this sentence? And of course the formal statistical analysis to go with it. Finally, as noted in the Report, OECD guidelines, IARC guidelines, NTP guidelines and others, the concurrent control is the best to use for evaluating a study.

Page 69 – "*This result was confirmed by the chi-square test. Also for this comparison, the very low control incidence (0/51) should be taken into consideration.*" Again, where are the historical controls to support this statement?

Page 71 – "*It may be concluded that the statistical significance of the suspected increase in malignant lymphoma in the various studies depends very much on the statistical method that is used for data analysis.*" This is usually the case; that is why the OECD guidelines make it clear that if either the trend test or the pairwise comparison is positive, the findings should be considered positive.

Page 71 – "*When the trend test is applied, the studies by Wood et al. (2009, ASB2012-11490) and Sugimoto (1997, ASB2012-11493) provide evidence of an effect which was not the case when pairwise comparison was performed. In contrast, the increase in the study of Kumar (2001, ASB2012-11491) was not confirmed neither by the trend test nor by a different pairwise test than the Z-test that had been used first.*" From my Table 6, there are two significantly positive studies, two studies with a marginal p-value and one study that would be positive if not for the highest dose dropping down. As noted in the Report, there was a drop in weight gain in the 1993 which could explain the drop in tumors at the highest exposure group (animals with reduced caloric intake are less likely to get tumors).

Page 71 – "*In the studies by Wood et al. (2009, ASB2012-11490) and by Atkinson et al. (1993, TOX9552382) in CD-1 mice, comparable top doses of 810 or 1000 mg/kg bw/day were administered and a similar incidence of malignant lymphoma was noted in high dose males (5/51 or 6/50, respectively). However, the control group incidences were clearly different (0/51 vs. 4/50) resulting in a positive trend test in the study by Wood et al. (2009, ASB2012-11490) only.*" The 1993 study was 24 months whereas the 2009 study was 18 months; it is not surprising the control tumor counts are higher in the 1993 study. What is surprising (and statistically significant) are the 6 tumors at the high dose in the 2009 study after only 18 months. And of course, this is another inappropriate comparison of control incidence over a 16 year timeframe. And finally, none of this is statistically significant.

Page 71 – "*Thus, if all four studies in CD-1 mice are taken together, there is no consistent dose response.*" See my formal analysis of this question.

Page 71 – "*Nonetheless, it seems well in line with information that was found in the literature providing confirmation that Swiss mice are prone to developing lymphoreticular tumours. According to older articles, control incidences in male mice of Swiss or Swiss-derived strains may reach 18–27.5% and exceed 36% in females (Sher, 1974, Z22020; Roe and Tucker, 1974, ASB2015-2534; Tucker, 1979, Z83266). In a more recent publication, Tadesse-Heath et al. (2000, ASB2015-2535) even mentioned a nearly 50% lymphoma (mostly of B cell origin) incidence in a colony of CFW Swiss mice but also emphasised the contribution of widespread infections with murine oncogenic viruses to the high but remarkably variable incidence of tumours of the lymphoreticular system in this species.*" Why are there guidelines if they are not used? Again, an argument is being made about historical controls using data which does not match OECD guidance (even bringing in Swiss-derived strains). And, if there are had historical control values from the lab, giving all five numbers and

some description of the studies (18 months or 24 months?) would seem to be in order.

Page 72 – "*However, in the study report itself, there was no evidence of health deterioration due to suspected viral infection and, thus, the actual basis of EPA's decision is not known.*" The entire discussion about infections is, at best, absurd if there is no evidence. Inclusion of this text is simply an attempt to discredit the study.

Page 72 – "*It ranged from 3.85% to 19.23% in the control groups from 12 studies that had been performed between 1992 and 1998 (Kitazawa, 2013, ASB2014-9146). Thus, the 12% incidence at the top dose level in the study with glyphosate was well covered by the range even though it was above the mean value of 6.33%.*" 12 studies with a mean of 6.33% and a range of 3.85 to 19.23 is an extremely skewed population. One study had 3.85% and one had 19.23 %; 12 x 6.33%=75.96 so the remaining 10 studies, in order to get an average of 6.33% would need to add up to 52.54 or 5.25% per study on average. Just from the math, it appears the 19.23% control is an outlier. Regardless, for sake of transparency, the actual rates should be given and assurances be given that they are all from studies of 18 months and not 24 months. And finally, a formal statistical analysis against the historical controls should be conducted. To illustrate; if the historical background is 6.33% and is based upon 50 animals in each control group and the controls are binomially distributed, then the probability of randomly seeing an outcome with a trend statistic equal to or larger than the one observed in this study is p=0.02. MATLAB code is provided that makes this calculation.

Page 72 – "*Unfortunately, for the study of Wood et al. (2009, ASB2012-11492), the submitted historical control data was not particularly useful for the assessment.*" Stop with this statement; everything else written is an inappropriate use of historical control data and should be ignored.

Page 73 – "*On balance, based on uncertainties with regard to partly contradictory study outcomes depending on the statistical method applied, inconsistent dose response in the individual studies, and a highly variable tumour incidence as suggested by historical control data, it is not likely that glyphosate has induced malignant lymphoma in mice. A possible role of oncogenic viruses should not be ignored. Moreover, human relevance of such an effect, if occurring only as a high-dose phenomenon as it was the case here, is considered equivocal.*" On balance, this entire paragraph is a wrong. The study outcomes are not contradictory (follow OECD guidance and it is simple), does response is not inconsistent (see my analysis), tumor incidence is not highly variable when properly adjusted for time on study differences and the entire historical control discussion is either inappropriate or inadequately applied.

Page 74 -"*Even though no historical control data from the performing laboratories was provided, a simple comparison of the control groups in the individual studies with*

*glyphosate suggests that renal tumours may occur in untreated control males at a similar incidence than in the groups receiving very high doses.*" This is a misleading comment. First, no formal analysis of historical control data has been undertaken and, as we stated in our paper (Portier *et al*., 2016), your own guidelines provide guidance on how to obtain and use historical control data; this has not been done here. I am also surprised to see the statement that "*no historical control data from the performing laboratories was provided*" when in response to a letter sent to Commissioner Andriukaitis, the EFSA Executive Director, Professor Url, wrote "*The Peer Review Report (EFSA, 2015b) confirms that EFSA conducted a specific check regarding the use of historical control data, requested additional information during the clock-stop procedure and only considered valid the historical control data from the performing laboratory in line with the international recommendations*". Which is it? Does the Report rely on valid historical control data from the performing laboratories or not?

Page 75 - "*Even if not fully comparable because of the strain differences, it should be remembered that the top dose incidence of 2/50 in this study was the same as seen in CD-1 mice in the study by Atkinson et al. (1993, TOX9552382) in the control and low dose groups.*" Why even include this sentence? They are not comparable.

Page 76 – "*Even though there was no clear dose response, it may be assumed that glyphosate (acid) when administered at high doses might produce mucosal irritation.*" So, if I am reading this right, statistically significant positive cancer results are being dismissed based on non-statistically significant non-cancer results that have a questionable linkage to the cancer results. Does this seem reasonable? I guess not since this appears in the next papagraph "*However, it is questionable if irritation would sufficiently explain tumour formation in the kidney.*".

Page 76 – "*The top dose finding of 2/50 in the study by Sugimoto (1997, ASB2012-11493) is at the upper edge of adenoma frequency. In the study by Knezevich and Hogan (1983, TOX9552381) which is not actually covered by the timeframe of the historical database, the adenoma incidence (2%) at the top dose level would be inside the historical range whereas a carcinoma incidence of 4% was above.*" Again, an improper use of historical controls. These are not appropriate for the 1983 study but are used anyway. For the 1997 study, only controls in mice sacrificed at 18 months should be used, mice sacrificed at 24 months will likely have greater incidence. This is quite evident when one looks at hemangiosarcomas in male mice in the Giknis and Clifford report (attached). Exactly half of the studies went 18 months, 24 went 2 years and the remaining two went 97 and 100 weeks. Hence this historical control dataset is inappropriate for this comparison. However, even if it were, the findings would still be significant. The paper gives a mean background level for adenomas of 0.24% and for adenocarcinoma of 0.14% for a combined background of 0.38%. The probability of seeing a dose-response trend equal to or larger than what was seen in the 1997 study is 0.01, a significant finding. The p-value for the 1983 study would be even smaller.

Page 77 – "*Even the incidences of affected animals at exaggerated doses exceeding the OECD-recommended limit of 1000 mg/kg bw/day and also the MTD were not statistically significantly increased when compared with the concurrent controls.*" As mentioned earlier in this document, if either test is positive, the findings should be considered positive so the second half of this sentence is inappropriate. How did "*there is some evidence that the MTD was exceeded in both studies at the highest dose level*" (Page 76) become absolute certainty about exceeding the MTD?

Page 77 – "*Even the incidences at exaggerated doses are covered by the historical control range.*" As noted earlier, this finding is not supported.

Page 77 – "*No pre-neoplastic kidney lesions have been observed in treated animals.*" Following this logic, the high dose animals got tumors by some unknown mechanism related to exceeding the MTD and that unknown mechanism did not damage the kidneys in any other animals enough to show preneoplastic effects. What is this mechanism and where is the evidence suggesting such a mechanism exists? And how does this statement "*However, it is questionable if irritation would sufficiently explain tumour formation in the kidney.*" fit in to this theory?

Page 77 – "*There is no plausible mechanism*" Following the logic again, some unknown mechanism related to exceeding the MTD caused the tumors at the highest doses and because there is no mechanism, the results should be dismissed.

Page 78 – "*According to Atkinson et al. (1993, TOX9552382), the historical control incidence in the performing laboratory ranged from 0/50 to 4/50 and, thus, would cover the incidence at the top dose level.*" Inadequate documentation of the historical control data makes it impossible to address this statement. The actual counts and ages at terminal sacrifice for the historical controls should be provided. As shown earlier, range is an inappropriate way to utilize historical controls. This is a clear example of a lack of transparency.

Page 78 – "*Historical control data provided by Charles River indicate a very variable incidence of haemangiosarcoma. On different sites of the body, tumours of this type were seen in untreated control animals in 8 of 52 studies.*" In this case, Giknis and Clifford give the actual values for each of their control groups. For hemangiosarcomas, there were zero tumors in all 26 studies terminated at 18 months, and only 8 of the remaining 26 studies that went two years had hemangiosarcomas. Thus, the 18 month 1997 study is well outside the range of the historical controls.

Page 78 – "*Furthermore, since Sugimoto (1997, ASB2012-11493) employed a more than four times higher top dose than Atkinson et al. (1993, TOX9552382), a markedly higher haemangiosarcoma incidence would have been expected if this tumour was in fact treatment-related.*" Again, this is a comparison of an 18 month study to a 24 month study. The finding that the 24 month 1993 study has an 8% response at a dose of 1000 mg/kg/day while the 18 month 1997 study has a 4% response at a 4-

fold higher dose is not unexpected.

# Table 1: Human Epidemiology Studies

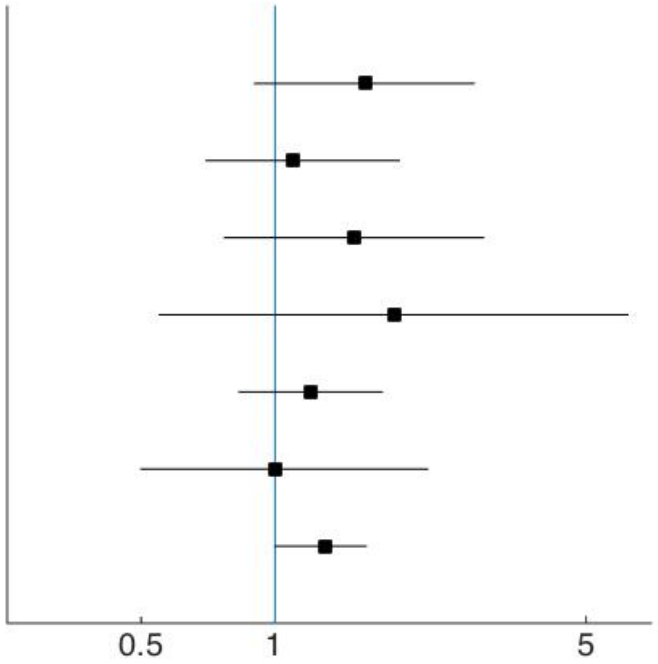
| Study | Type | Size | Findings | Exposed Cases |
|---|---|---|---|---|
| Agricultural Health Study *(De Rooset al., 2005)* | Cohort – licensed pesticide applicators | 52 395 (+32 347 spouses), 92 cases, 4-8 years follow-up | 1.1 (0.7-1.9) C<br>0.7 (0.4-1.4) 21-56% tertile compared to <20% tertile<br>0.9 (0.5-1.6) 21-56% tertile compared to >57% tertile<br>(31 cases no quantification of exposure) | 73 |
| US Midwest *(De Roos et al., 2003)* | Pooled analysis 3 case-control studies | NHL: 650 cases, 1933 controls | 2.1 (1.1-4) U<br>1.6 (0.9-2.8) C | 36<br>36 |
| Cross-Canada *(McDuffie et al., 2001)* | Population-based case-control study | 517 cases, 1506 controls | 1.2 (0.83-1.74) U<br>1.0 (0.63-1.57) ≤2 d/Y<br>2.12 (1.2-3.73) >2 d/Y | 51<br>28<br>23 |
| Swedish Case-Control Study *(Eriksson et al., 2008)* | Population-based case-control study | 910 cases, 1016 control | 2.02 (1.1-3.71) U<br>1.51 (0.77-2.94) C<br>1.69 (0.7-4.07) ≤10 d/Y<br>2.36 (1.04-5.37) >10 d/Y<br>1.11 (0.24-5.08) ≤10 Y<br>2.26 (1.16-4.4) >10 Y | 29<br>29<br>12<br>17<br>NR<br>NR |
| Swedish Case-Control Study *(Hardell et al., 1999)* | Population-based case-control study | 404 cases, 741 control (limited power) | 2.3 (0.4-1.3) U<br>5.8 (0.6-5.4) C (not specified) | 4<br>NR |
| France Case-Control *(Orsi et al, 2009)* | Hospital-based case-control study | 244 cases, 456 controls | 1.0 (0.5-2.2) U | 12 |
| Swedish Case-Control Study *(Hardell et al., 2002)* | Population-based case-control study | 515 cases, 1141 controls | 3.04 (1.08-8.5) U<br>1.85 (0.55-6.2) C (notspecified) | 8<br>8 |
| US Case-Control Study *(Lee et al., 2004)* | Population-based case-control study | 872 cases, 2381controls | 1.4 (0.98-2.1) U – no asthma<br>1.2 (0.4-3.3) U - asthma | 53<br>6 |

# Table 2: Meta Analyses

| Study | Included Studies | Findings |
|---|---|---|
| Schinasi and Leon, 2014 | McDuffie et al., 2001; Hardell et al., 2002; De Roos et al., 2003 and 2005; Eriksson et al., 2008; Orsi et al., 2009) | 1.5 (1.1-2.0) |
| IARC Monograph Working Group | McDuffie et al., 2001; Hardell et al., 2002; De Roos et al., 2003 and 2005; Eriksson et al., 2008; Orsi et al., 2009) | 1.3 (1.103-1.65) – used adjusted risk estimates from Hardell et al., 2003 and Eriksson et al., 2008 |
| Chang and Delzell, 2016 | McDuffie et al., 2001; Hardell et al., 2002; De Roos et al., 2003 and 2005; Eriksson et al., 2008; Orsi et al., 2009) | 1.3 (1.0-1.6) |

Figure 1: Tree Plot of Epidemiology Studies

| Study | RR | Lower | Upper | Weight |
|---|---|---|---|---|
| De Roos et al. (2003) | 1.600 | 0.900 | 2.800 | 16.2 |
| De Roos et al. (2003) | 1.100 | 0.700 | 1.900 | 21.0 |
| Eriksson et al., (2008) | 1.510 | 0.770 | 2.940 | 11.6 |
| Hardell et al. (2002) | 1.850 | 0.550 | 6.200 | 3.6 |
| McDuffie et al. (2001) | 1.200 | 0.830 | 1.740 | 38.1 |
| Oris et al. (2009) | 1.000 | 0.500 | 2.200 | 9.5 |
| Meta-Analysis | 1.300 | 1.000 | 1.600 | 100.0 |




Chang and Delzell (2016)

# Table 3: Carcinogenicity Studies in Male Mice

| Year | Strain | Length[1] | Top Dose[2] | Renal Tumors | Hemangio-sarcomas | Malignant Lymphoma |
|---|---|---|---|---|---|---|
| 1983[5] | Crl:CD-1 | 24 | 4,841 | +[3] | | |
| 1993[5] | ?:CD-1 | 24 | 1,000 | | + | +/-[4] |
| 1997 | CrJ:CD-1 | 18 | 4,843 | + | + | + |
| 2001 | SW | 18 | 1,460 | + | **Data Not Available** | + |
| 2009 | Crl:CD-1 | 18 | 810 | | | + |

1 – months; 2 – mg/kg bw/day; 3 - + indicates a p-value of <0.05 as calculated by BfR using the Armitage linear trend test in proportions; 4 – p=0.054; 5 – studies evaluated in IARC review; p=0.08

# Table 4: Analysis of Male Mouse Renal Tumors From the Individual Studies

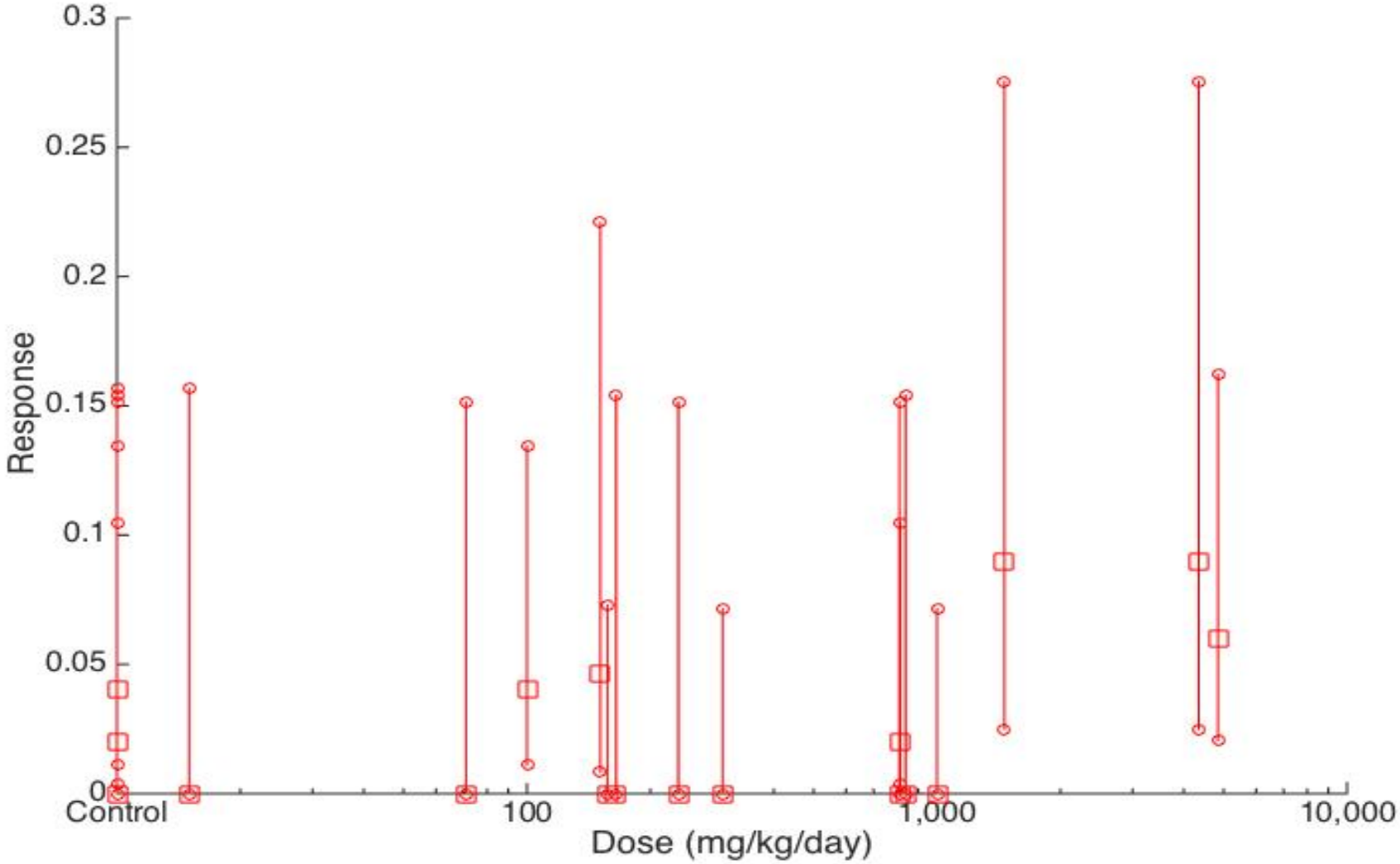
| Year | Strain | Length | Doses (mg/kg/d) | Response | p-Trend (p-poly3) |
|---|---|---|---|---|---|
| 1983 | Crl:CD-1 | 24 | 157, 814, 4841 | 1/50, 0/49, 1/50, 3/50 | 0.03 (0.03) |
| 1993 | ?:CD-1 | 24 | 100, 300, 1000 | 2/50, 2/50, 0/50, 0/50 | 0.94 (0.94) |
| 1997 | CrJ:CD-1 | 18 | 165, 838, 4348 | 0/50, 0/50, 0/50, 2/50 | 0.008 (0.009) |
| 2001 | SW | 18 | 15, 151, 1460 | 0/49, 0/49, 1/50, 2/50 | 0.04 (0.04) |
| 2009 | Crl:CD-1 | 18 | 71, 234, 810 | 0/51, 0/51, 0/51, 0/51 | - |

# Table 5: Pooled Analysis of Male Mouse Renal Tumors

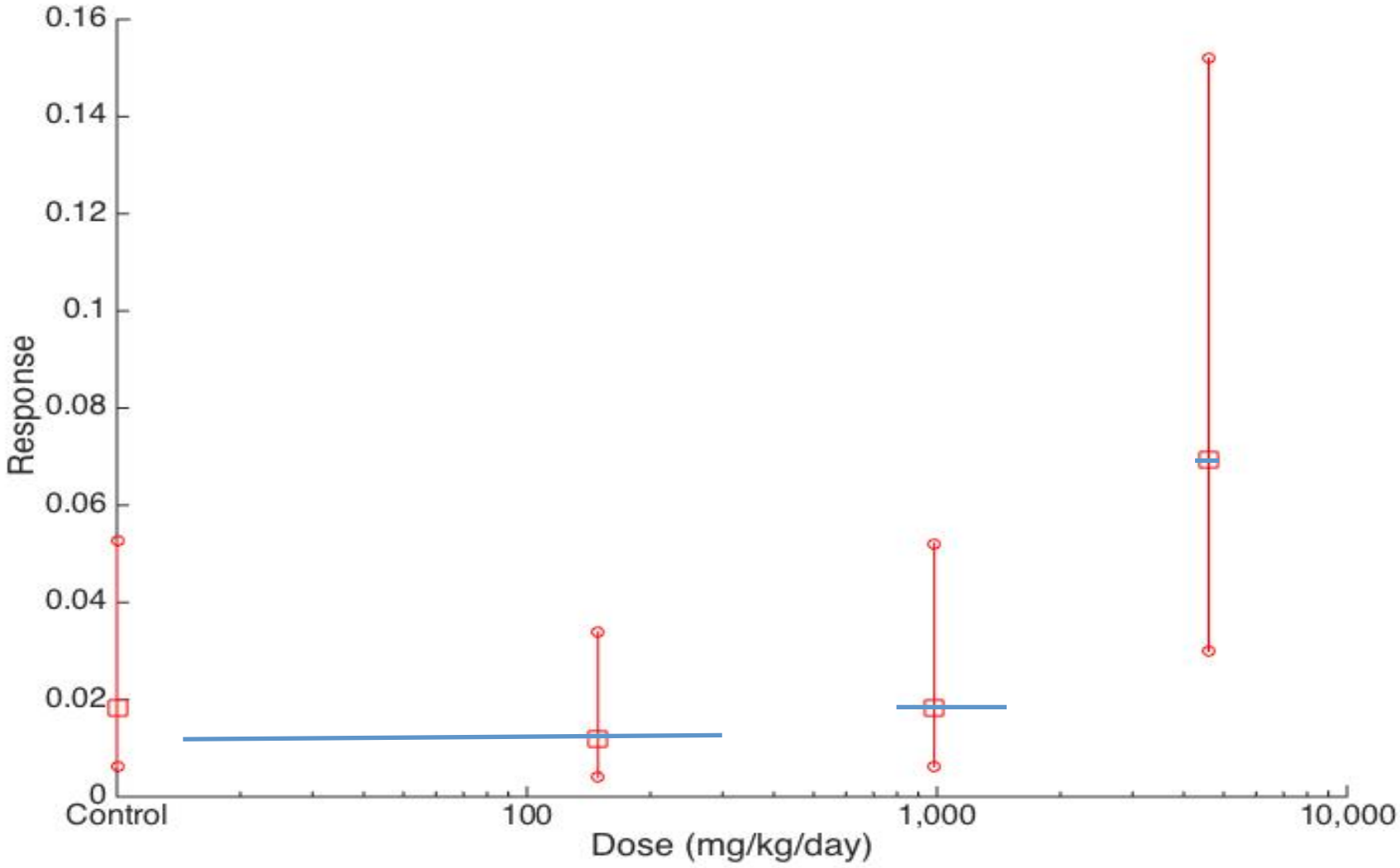
| Year | Strain | p-Trend (p-poly3) |
|---|---|---|
| All Combined | CD-1 and Swiss | 0.0004 (0.001) |
| All Combined and Doses Pooled[1] | CD-1 and Swiss | 0.0008 (0.002) |
| CD-1 Combined | CD-1 | 0.001 (0.001) |
| CD-1 Combined and Doses Pooled[1] | CD-1 | 0.001(0.001) |
| All Combined, doses>1000 dropped | CD-1 and Swiss | 0.80 (0.84) |
| All Combined, doses>1000 dropped and Doses Pooled[2] | CD-1 and Swiss | 0.39 (0.40) |
| CD-1 Combined, doses>1000 dropped | CD-1 | 0.85 (0.86) |
| CD-1 Combined, doses>1000 dropped and Doses Pooled[2] | CD-1 | 0.80 (0.80) |

[1] – Doses were combined as follows: all controls, doses between 0 and 310 mg/kg/day, doses between 310 and 1500 mg/kg/day, and doses greater than 1500 mg/kg/day. Average doses in each pooled group were used in the analysis. [2-] Doses were combined as follows: all controls, doses between 0 and 310 mg/kg/day, and doses between 310 and 1500 mg/kg/day. Average doses in each pooled group were used in the analysis.

# Figure 2: Renal tumors in male mice poly-3 adjusted showing individual dose groups



# Figure 3: Renal tumors in male mice poly-3 adjusted and clustered by similar doses



# Table 6: Analysis of Male Mouse Malignant Lymphoma From the Individual Studies

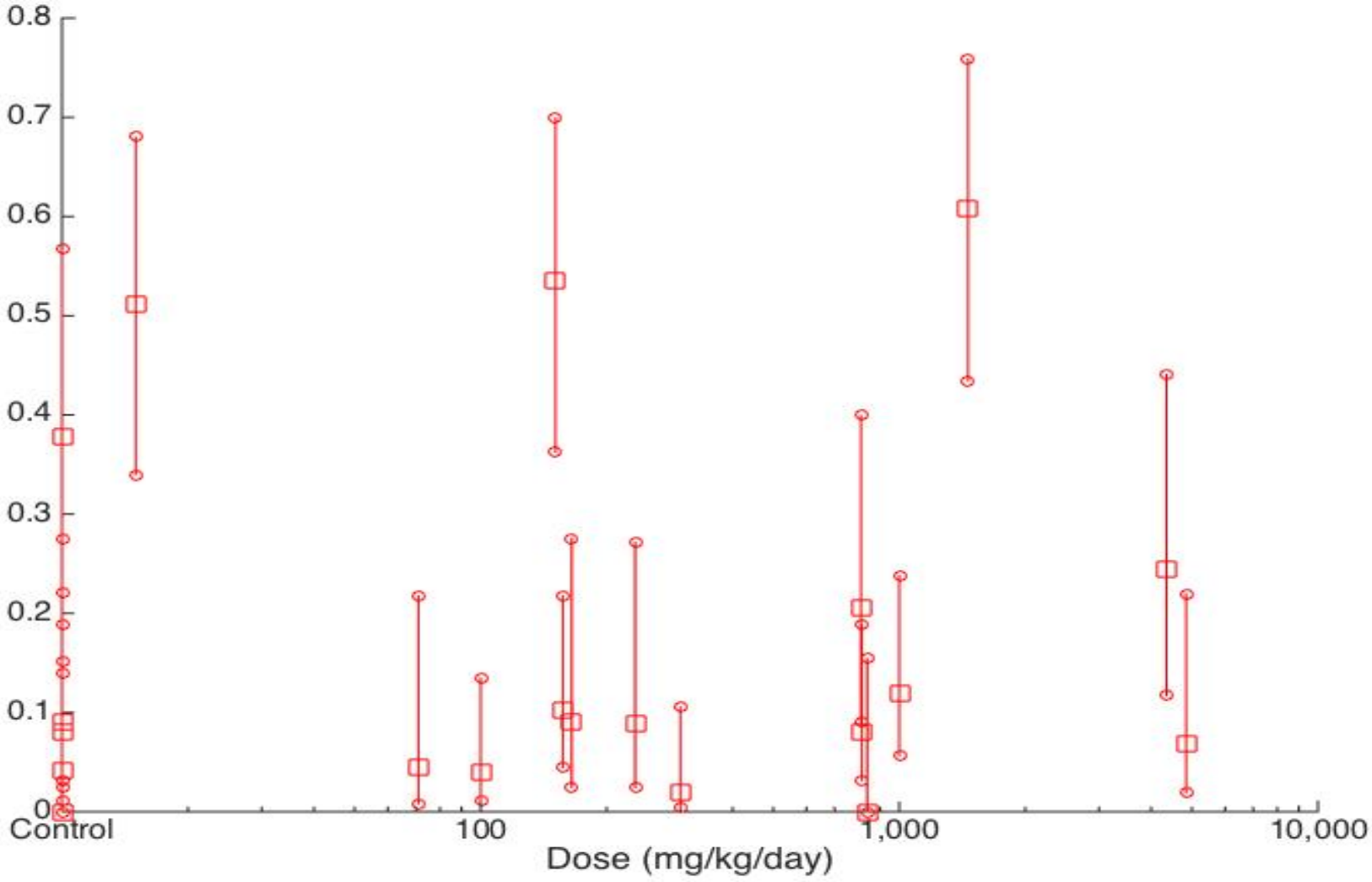
| Year | Strain | Length | Doses (mg/kg/d) | Response | p-Trend (p-poly3) |
|---|---|---|---|---|---|
| 1983 | Crl:CD-1 | 24 | 157, 814, 4841 | 2/50, 5/49, 4/50, 2/50 | 0.51 (0.51) |
| 1993 | ?:CD-1 | 24 | 100, 300, 1000 | 4/50, 2/50, 1/50, 6/50 | 0.08 (0.08) |
| 1997 | CrJ:CD-1 | 18 | 165, 838, 4348 | 2/50, 2/50, 0/50, 6/50 | 0.008 (0.012) |
| 2001 | SW | 18 | 15, 151, 1460 | 10/49, 15/49, 16/49, 19/49 | 0.05 (0.09) |
| 2009 | Crl:CD-1 | 18 | 71, 234, 810 | 0/51, 1/51, 2/51, 5/51 | 0.004 (0.005) |

# Table 7: Pooled Analysis of Male Mouse Malignant Lymphoma

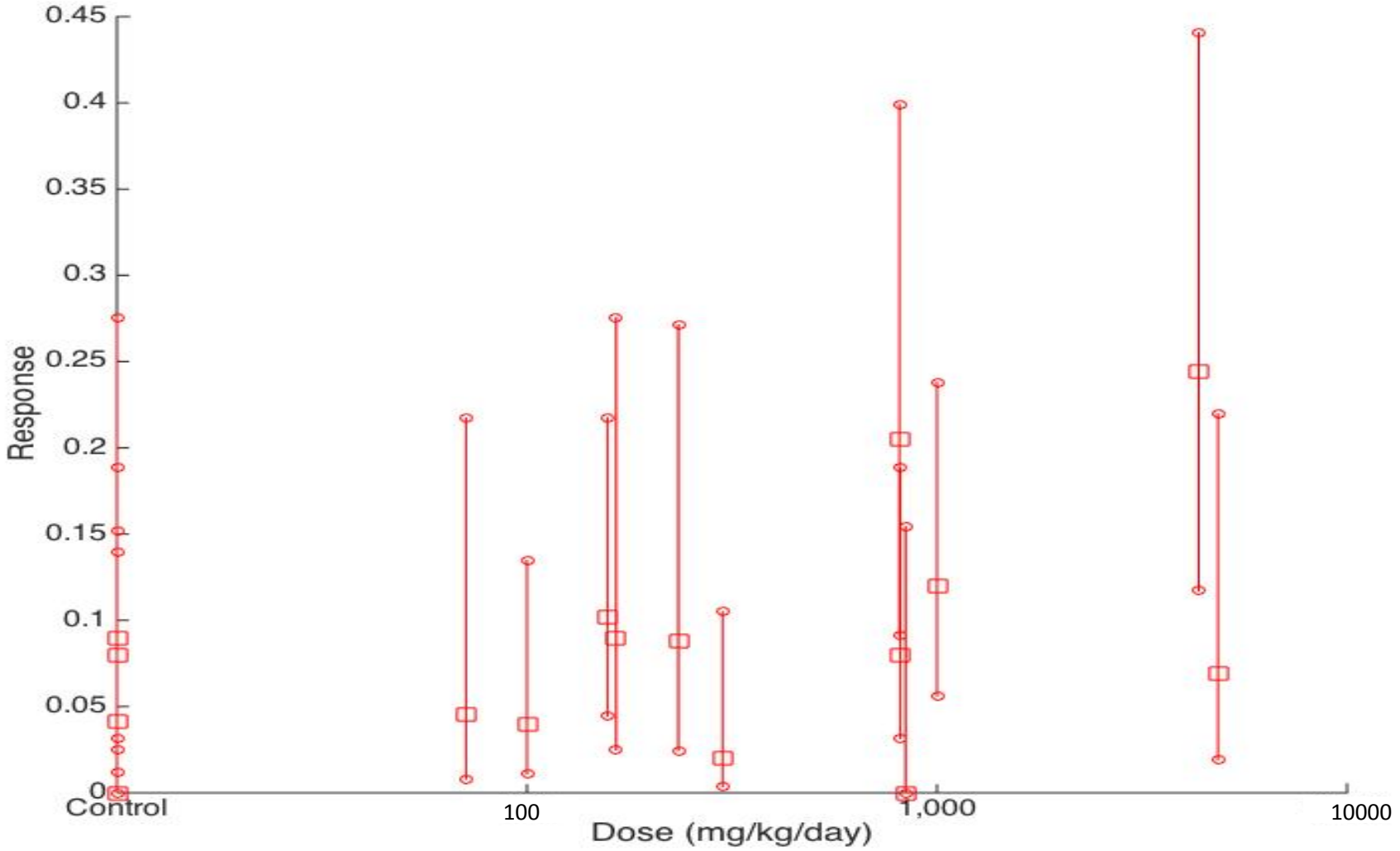
| Year | Strain | p-Trend (p-poly3) |
|---|---|---|
| All Combined | CD-1 and Swiss | 0.17 (0.19) |
| All Combined and Doses Pooled[1] | CD-1 and Swiss | 0.32 (0.31) |
| CD-1 Combined | CD-1 | 0.02 (0.01) |
| CD-1 Combined and Doses Pooled[1] | CD-1 | 0.01(0.009) |
| All Combined, doses>1000 dropped | CD-1 and Swiss | 0.86 (0.93) |
| All Combined, doses>1000 dropped and Doses Pooled[2] | CD-1 and Swiss | 0.02 (0.03) |
| CD-1 Combined, doses>1000 dropped | CD-1 | 0.03 (0.05) |
| CD-1 Combined, doses>1000 dropped and Doses Pooled[2] | CD-1 | 0.04 (0.04) |

[1] – Doses were combined as follows: all controls, doses between 0 and 310 mg/kg/day, doses between 310 and 1500 mg/kg/day, and doses greater than 1500 mg/kg/day. Average doses in each pooled group were used in the analysis. [2]- Doses were combined as follows: all controls, doses between 0 and 310 mg/kg/day, and doses between 310 and 1500 mg/kg/day. Average doses in each pooled group were used in the analysis.

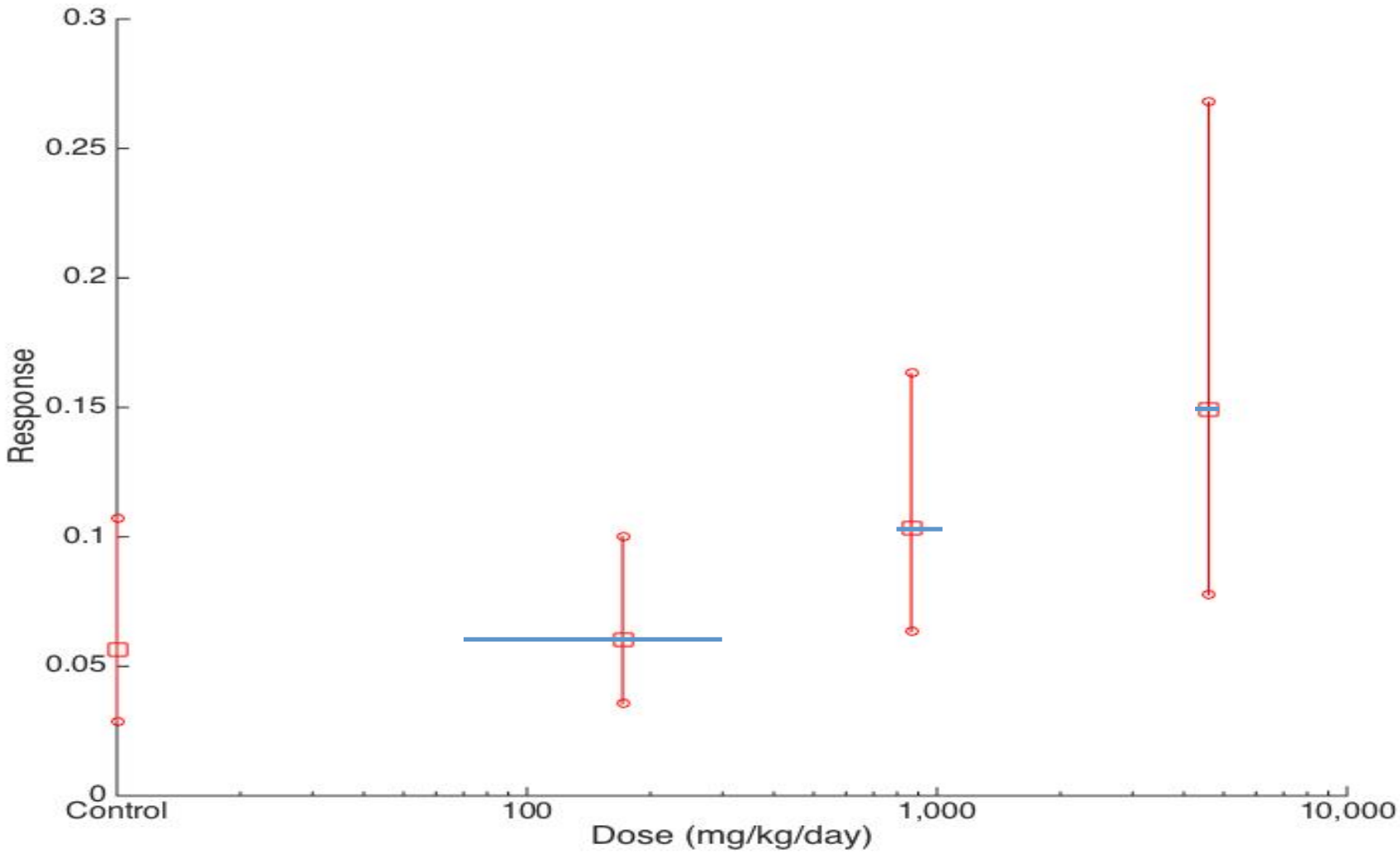
# Figure 4: Malignant lymphomas in male mice poly-3 adjusted showing individual dose groups



# Figure 5: Malignant lymphomas in male CD-1 mice poly-3 adjusted showing individual dose groups



# Figure 6: Malignant lymphomas in male CD-1mice poly-3 adjusted and clustered by similar doses



# Table 8: Analysis of Male Mouse Hemangiosarcomas From the Individual Studies

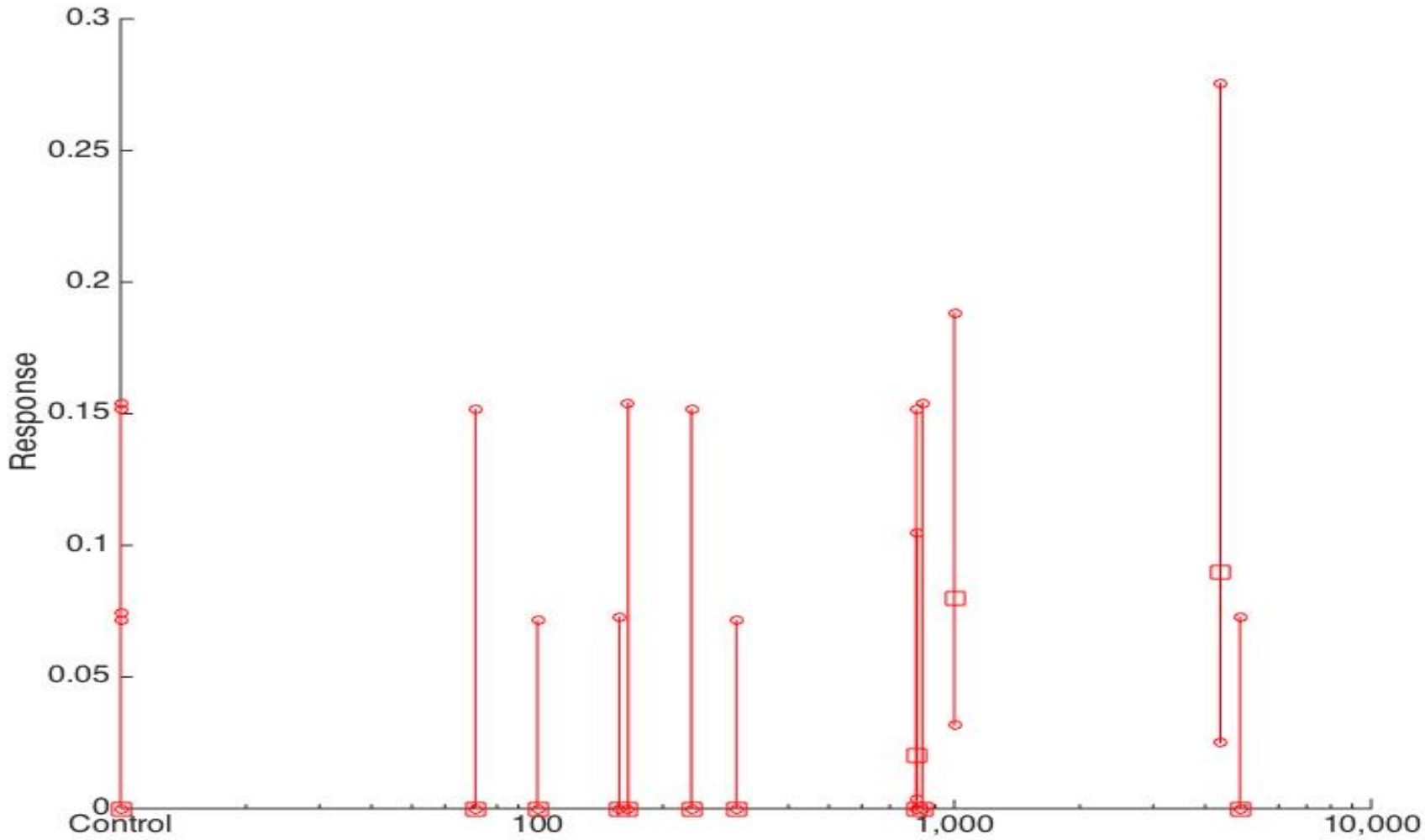
| Year | Strain | Length | Doses (mg/kg/d) | Response | p-Trend (p-poly3) |
|---|---|---|---|---|---|
| 1983 | Crl:CD-1 | 24 | 157, 814, 4841 | 0/50, 0/49, 1/50, 0/50 | 0.63 (0.63) |
| 1993 | ?:CD-1 | 24 | 100, 300, 1000 | 0/50, 0/50, 0/50, 4/50 | 0.0004 (0.0004) |
| 1997 | CrJ:CD-1 | 18 | 165, 838, 4348 | 0/50, 0/50, 0/50, 2/50 | 0.008 (0.009) |
| 2001 | SW | 18 | 15, 151, 1460 | No Data | - |
| 2009 | Crl:CD-1 | 18 | 71, 234, 810 | 0/51, 0/51, 0/51, 0/51 | - |

# Table 9: Pooled Analysis of Male Mouse Hemangiosarcomas

| Year | Strain | p-Trend (p-poly3) |
|---|---|---|
| CD-1 Combined | CD-1 | 0.02 (0.03) |
| CD-1 Combined and Doses Pooled[1] | CD-1 | 0.02 (0.02) |
| CD-1 Combined, doses>1000 dropped | CD-1 | <0.0001 (<0.0001) |
| CD-1 Combined, doses>1000 dropped and Doses Pooled[2] | CD-1 | 0.0003 (0.0003) |

[1] – Doses were combined as follows: all controls, doses between 0 and 310 mg/kg/day, doses between 310 and 1500 mg/kg/day, and doses greater than 1500 mg/kg/day. Average doses in each pooled group were used in the analysis. [2-] Doses were combined as follows: all controls, doses between 0 and 310 mg/kg/day, and doses between 310 and 1500 mg/kg/day. Average doses in each pooled group were used in the analysis.

# Figure 7: Hemangiomas in male CD-1 mice poly-3 adjusted showing individual dose groups



# Figure 8: Hemangiomas in male CD-1 mice poly-3 adjusted and clustered by similar doses